Final Project Proposal

CUNY 620

Group 4 – Niteen Kumar, Alexander Low, Jagruti Solao, Brian Weinfeld

For our final project we wish to examine the sentiment of social media posts based on the live results of the World Cup. We are interested in exploring how the sentiments of tweets change as a teams likelihood of success rises and falls. Do fans become more negative after their team is eliminated? Are they happy with their team’s run? Can we identify key moments in the tournament or individual games via the change in sentiment? Is it feasible to expand this outside the world of sports?

**Work Flow:**

We will write code to scrape Twitter for tweets that have various hashtags. From there we will perform sentiment analysis on the tweet in order to determine how positive or negative the tweet is. Using the timestamp for the tweet we will attempt to align the tweet in a timeline corresponding to World Cup events. We may also use tweet volume as a metric.

We plan on starting with a small focus on an individual team (England) and then will broaden our scope based on our initial results.

After collecting our results we will attempt to answer our initial guiding questions. In totality we will attempt to determine whether Twitter can be used to judge the general fortunes of specific events in near real time.

**Concerns:**

There are a number of concerns about our ability to turn the raw data into actionable information. We must search for tweets via hashtag so we must carefully select the hashtags that we wish to have represent our data set. We must also be able to perform sentiment analysis on text that is informal and slangy. We must also be careful about trying to identify key moments in the World Cup and changes in sentiment in the tweets we scrape.

**Roles:**

Our group has no set, single defined rolls. Instead we plan to work closely together, either simultaneously working on different parts of the project or handing off the project to each member in turn order.